How Physical Measurements Really Work: Forward Models, Inverse Inference, Calibration & Systematics

Ervin Kafexhiu

Department of Physics
Faculty of Natural Sciences, University of Tirana

ICMMDSRI 2025

16 December 2025

Why this lecture?

Many students (and PhDs!) can take data but struggle to:

- **explain what was actually measured**,
- **identify assumptions,**
- **reason about systematic errors.**

Most experimental failures are not technical — they are conceptual.

Goal of this lecture:

To provide a general way of thinking about measurements, independent of the specific instrument.

Why measurements fail

What many students implicitly assume

Instrument → **produce data** → **fit a model** → **get a result** ⇢ **Discovery?**

Everything here is true — and still completely misleading.

What is actually happening: **Instrument = hardware + software + assumptions**

- You made **many assumptions** before seeing any data.
- Some assumptions are **implicit or forgotten**.
- **Your result is valid only within those assumptions.**

If your assumptions are wrong or violated,
the measurement is meaningless — even if the data look perfect.

Most of you have produced beautiful plots that mean absolutely nothing.

Not because you are careless — but because assumptions were never made explicit!

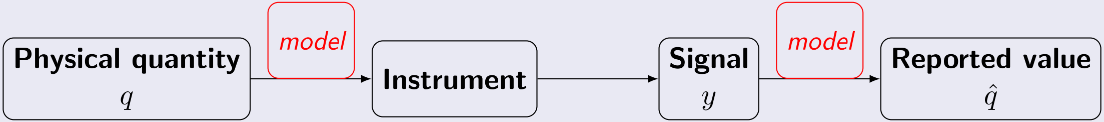
Key message

Do not forget

A measurement is not a number.

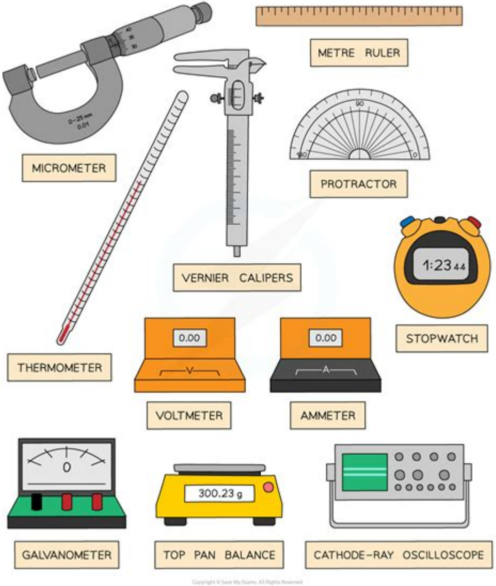
A measurement is an inference based on a model.

From physical quantity to reported result



- The instrument does not report q directly.
- Arrows are models that *claim* to connect the world to the reported value.
- Every arrow hides assumptions.
- Understanding these arrows is understanding measurement.

This is the entire logic of measurement. Everything we discuss next is about these arrows.



METRE RULER
MICROMETER
PROTRACTOR
VERNIER CALIPERS
STOPWATCH
1:23 44
THERMOMETER
0.00
VOLTMETER
0.00
AMMETER
0
GALVANOMETER
300.23 g
TOP PAN BALANCE
CATHODE-RAY OSCILLOSCOPE

Two different questions are hidden in every measurement

Every measurement implicitly answers *two different questions*.

1. **If the physical world is like this, what signal does my instrument produce?**
2. **Given the signal I measured, what can I say about the physical world?**

These two questions are not the same.
Confusing them is a source for measurement failures.

Forward & Inverse Models

Forward model: from the world to the signal

Definition

The **forward model** describes how a physical state of the world produces the signal measured by the instrument.

- It is **predictive**: world $\rightarrow$ signal.
- It is determined by:
  - physics of the interaction,
  - geometry and timing,
  - instrument response,
  - environmental conditions.
- It exists **before any data are taken**.

Key idea

If you cannot clearly state your forward model, you do not know what your instrument is sensitive to.

Forward model: conceptual functional form

Conceptual form

$$y = f(q, \theta) + \varepsilon$$

- q : physical quantity (or state of the world)
- θ : instrument and environmental parameters
- ε : noise and unmodeled effects

Interpretation

The forward model predicts what signal the instrument *would produce* for a given physical state.

Forward model: key properties

- **Many-to-one mapping**

$$f(q_1, \theta) \approx f(q_2, \theta)$$

Different physical states can produce the same signal.

- **Noise**

$$f(q, \theta) + \varepsilon \neq f(q, \theta)$$

The same physical state never produces exactly the same signal twice.

- **Approximate model**
  - unknown parameters,
  - neglected physics,
  - uncontrolled environment.

Consequence

Even if physics is deterministic, the forward model is not exactly invertible in practice.

Inverse model: from signal to inference

Definition

The **inverse model** describes how we infer a physical quantity
from the measured signal.

- It is an **inference**: signal $\rightarrow$ world.
- It requires:
  - assumptions,
  - constraints,
  - algorithms,
  - often simplifications or approximations.
- It is **not unique** in general. Different algorithms, different assumptions, and different priors can give different answers from the same data.

Key idea

The inverse model is where most assumptions hide.

Inverse model: conceptual functional form

Conceptual form

$$\hat{q} = g(y \mid \text{assumptions, priors, constraints})$$

- y : measured signal
- assumptions : physics, geometry, simplifications
- priors : what is considered likely or reasonable
- constraints : mathematical or physical restrictions

Interpretation

The inverse model expresses what we infer about the world given the data *and* our assumptions.

Inverse model: key properties

- **Not unique**
  - different algorithms,
  - different priors,
  - different constraints

  can **give different answers from the same data**.

- **Requires choices**
  - what to ignore,
  - what to regularize[1],
  - what to assume is likely.

- **Algorithm-dependent** Data analysis is the practical implementation of the inverse model.

[1] Regularization means adding assumptions that prevent the solution from doing unphysical or unstable things. If many solutions fit the data equally well, regularization is how we choose the simplest or most reasonable one.

Why the inverse model is not f^{-1}

Key point: In general, $g \neq f^{-1}$.

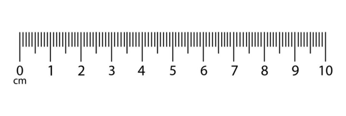
- **The forward model loses information.**
- **Noise breaks exact invertibility.**
- **Models are incomplete.**
- **Inference requires assumptions and choices.**

The forward model tells you what your instrument would measure.
The inverse model tells you what you are willing to believe given the data.

Example 1: Measuring length with a ruler

Forward model

Physical length $L \longrightarrow$ **visual alignment with scale marks**



Inverse model

Observed mark position $\longrightarrow$ **inferred length** $\hat{L}$

Key assumptions

- ruler is straight and uniformly spaced
- object is aligned with the ruler
- no parallax error

True length L $\xrightarrow{\text{forward}}$ Ruler $\longrightarrow$ Observed marks $\xrightarrow{\text{inverse}}$ Inferred length $\hat{L}$

Example 2: Liquid-in-glass thermometer

Forward model

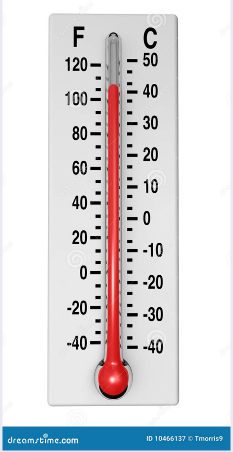
Temperature $T \longrightarrow$ **thermal expansion** $\longrightarrow$ **liquid height** $h(T)$

Inverse model

Measured height $h \longrightarrow$ **calibration curve** $\longrightarrow \hat{T}$

Key assumptions

- thermal equilibrium
- linear or known expansion law
- stable material properties



Example 3: Thermistor temperature sensor

Forward model

Temperature $T \longrightarrow$ **material physics** $R(T) \longrightarrow$ **voltage signal** V

Inverse model

Measured voltage $V \longrightarrow$ **circuit + calibration** $\longrightarrow \hat{T}$

Key assumptions

- known bias current or voltage
- negligible self-heating
- stable electronics

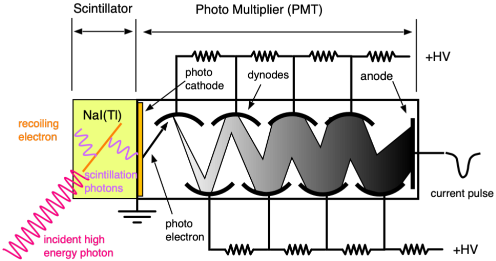
Example 4: NaI scintillation detector

Forward model

Particle energy E $\longrightarrow$ **energy deposition** $\longrightarrow$ **scintillation photons** $\longrightarrow$ **electronic signal**

Inverse model

Measured spectrum $\longrightarrow$ **statistical inference** $\longrightarrow$ $\hat{E}$ or flux



Key assumptions

- known scintillation yield
- stable PMT gain
- statistical photon production

What these examples have in common

- No instrument measures the quantity directly.
- Physics mediates between the world and the signal.
- Forward models can be analytic or simulated.
- Inverse models always require assumptions.

If you cannot draw this chain, you do not understand the measurement.

Calibration, Validation, Systematics

Calibration: what it really is

Definition

Calibration is the process of determining the unknown parameters of the forward model using known reference inputs.

- It does **not** define the model.
- It **assumes** the forward model is correct.
- **It assigns values to model parameters.**

Key idea

Calibration makes a model usable, not true.

Calibration inside the forward model

Forward model (recalled)

$$y = f(q, \theta) + \varepsilon$$

- q : physical quantity of interest
- θ : calibration parameters
- ε : noise and unmodeled effects

Calibration is the determination of θ using reference measurements with known q.

Calibration is not separate from the forward model. It lives inside it, in the parameters.

What calibration does *not* fix

- A wrong forward model
- Missing physics
- Invalid assumptions
- Operating outside the calibration regime
- Environmental changes after calibration

Key warning

A perfectly calibrated wrong model gives precise but meaningless results.

Calibration is the moment where our belief in the forward model meets reality.

Simple examples of calibration

- **Ruler**: scale markings define the length response.
- **Thermometer**: ice and boiling water define temperature mapping.
- **Detector**: known radioactive sources define energy scale.
- **Camera**: known distances define pixel-to-length conversion.

Common structure

Known input $\rightarrow$ measured signal $\rightarrow$ parameter estimation

Validation: what is it?

Definition

Validation is the process of testing whether the forward model and its assumptions describe reality beyond the calibration data.

- Calibration answers: *"What are the parameters?"*
- Validation answers: *"Is the model itself acceptable?"*

Key idea

Validation asks whether the model makes sense in situations it was not tuned for.

Validation is not fitting

- Using the same data to calibrate and validate is circular.
- A good fit does **not** imply a correct model.
- Validation must rely on:
  - independent data,
  - theoretical expectations,
  - symmetries or limiting cases,
  - alternative measurement methods.

Warning

A model can fit the data perfectly and still be wrong.

A good fit only tells you the model can reproduce the data — not that it explains it.

How do we validate a measurement?

- **Consistency checks**
  - repeat measurements,
  - stability over time.
- **Physical expectations**
  - known scaling laws,
  - conservation laws,
  - symmetry arguments.
- **Limiting cases**
  - extreme values,
  - simplified regimes.
- **Independent methods**
  - different instruments,
  - different models.

Residuals: where models reveal their limits

Residuals

$$r = y_{\text{measured}} - y_{\text{model}}$$

- Random residuals → model may be adequate.
- Structured residuals → missing physics or wrong assumptions.
- Drift or trends → instability or unmodeled effects.

Key insight

Residuals test the model, not the noise.

Residuals are not garbage. They are the model talking back to you!

Validation and systematic errors

- Failed validation often indicates:
  - broken assumptions,
  - missing physics,
  - unaccounted environmental effects.
- These failures are not random noise.

Key idea

Systematic errors are revealed by validation.

Calibration tells you what numbers to use;
validation tells you whether you should trust them.

Systematic errors: what are they?

Definition

Systematic errors are reproducible biases in a measurement caused by incorrect, incomplete, or violated model assumptions.

- They do **not** average out with more data.
- They shift results in a consistent direction.
- They originate from the model, not the noise.

Key idea

Systematic errors are model errors disguised as measurements. They originate from bad assumptions (not randomness in data).

Random vs systematic errors

- **Random errors**
  - fluctuate from measurement to measurement,
  - decrease with averaging,
  - are usually visible as scatter.
- **Systematic errors**
  - bias results consistently,
  - do not decrease with averaging,
  - are often invisible in error bars.

More data reduce noise. They do **not** fix systematic errors.

Where do systematic errors come from?

- **Forward model assumptions**
  - geometry,
  - linearity,
  - stationarity,
  - neglected physics.
- **Calibration assumptions**
  - reference validity,
  - stability over time,
  - operating regime.
- **Inverse model assumptions**
  - priors,
  - regularization,
  - algorithmic choices.

Every assumption we made earlier has the potential to become a systematic error.

How do systematic errors reveal themselves?

- Disagreement with known physical laws
- Inconsistent results across conditions
- Structured residuals
- Dependence on experimental setup
- Unexpected correlations

Key insight

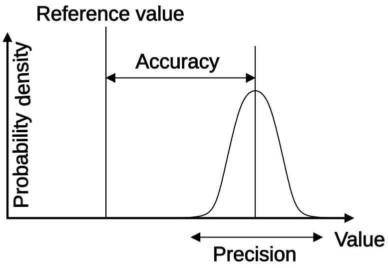
Systematic errors are revealed by validation, not by fitting.

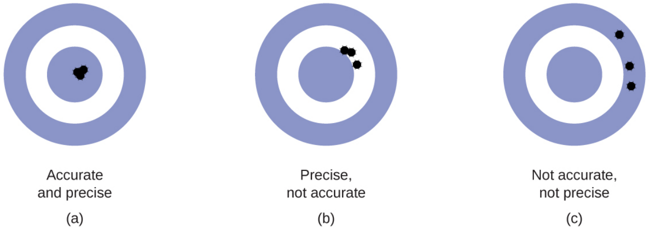
If something looks too clean, that's often a warning sign.

Why systematic errors are dangerous

- They produce precise but wrong results (not accurate).
- They survive averaging and statistics.
- They often go unnoticed without validation.

Small error bars do not imply correctness. They only imply consistency.





How do we deal with systematic errors?

- Make assumptions explicit
- Vary conditions intentionally
- Compare independent methods
- Stress-test models
- Revisit the forward model

Key idea

You do not eliminate systematic errors.
You identify, estimate, and control them.

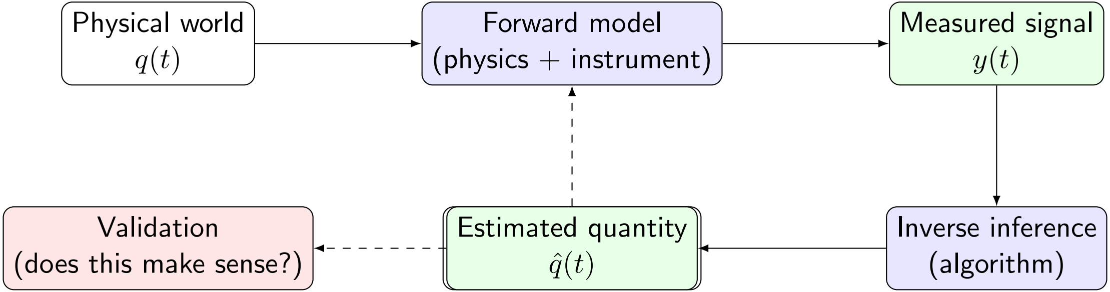
Random errors limit precision; systematic errors limit truth.

A universal workflow for physical measurements

1. **Define the measurand** What physical quantity do I claim to measure?
2. **Write the forward model** How does the world produce my signal?
3. **List assumptions** Geometry, linearity, stability, neglected physics.
4. **Calibrate the model** Determine unknown parameters using known references.
5. **Acquire data** Under controlled and documented conditions.
6. **Define the inverse model** How do I infer the measurand from the signal?
7. **Quantify uncertainty** Include noise, calibration, and model uncertainty.
8. **Validate** Test against independent expectations or limits.
9. **Identify systematic errors** Ask which assumptions could bias the result.

This list is not a recipe. It's a way of thinking.

You don't always do these steps once. You loop over them — especially when validation fails.



A measurement is not a number you obtain — it is an inference based on a model.

Thank You!